Identification of the porcine paramyxovirus LPMV matrix protein gene: comparative sequence analysis with other paramyxoviruses

Mikael Berg, Anita Sundqvist, Jorge Moreno-López and Tommy Linné*

Department of Veterinary Microbiology (Molecular Virology), Biomedical Center, Swedish University of Agricultural Sciences, Uppsala S-751 23, Sweden

The complete nucleotide sequence of the gene encoding the matrix protein (M) of the porcine paramyxovirus LPMV has been determined. The gene is 1376 nucleotides long including 5' and 3' non-coding sequences with a protein-coding sequence of 1107 nucleotides. The deduced protein, containing 369 amino acids with a calculated Mr of 41657, is hydrophobic overall with a net positive charge of +17.5. Comparative sequence analysis revealed high amino acid homology to other paramyxovirus M proteins, with the highest degree of identity (46%) with the human mumps virus. This is strong evidence that the porcine paramyxovirus LPMV is a genuine member of the paramyxovirus genus.

Introduction

La Piedad, Michoacan virus (LPMV) is a newly identified porcine paramyxovirus causing a fatal disease in piglets. The disease was characterized mainly by encephalomyelitis, pneumonia, reproductive failure (sows) and corneal opacity (blue eye) (Stephano et al., 1982, 1988; Moreno-López et al., 1986) and experimental infections demonstrated LPMV to be the causative agent (Moreno-López et al., 1986; Stephano et al., 1988). The clinical signs and the endemic in Mexico since the first outbreak in 1980 have been reviewed by Stephano et al., (1988).

Paramyxoviridae are divided into three genera: parainfluenzavirus, morbillivirus and pneumovirus (Matthews, 1982). LPMV was shown to have a protein composition and general features common to the paramyxovirus genera, i.e. haemagglutination, neuraminidase activity and syncytium formation of infected cells (Stephano et al., 1982, 1988; Moreno-López et al., 1986; Sundqvist et al., 1990). From the protein profile of the LPMV virion it was observed that the highly phosphorylated phosphoprotein (P protein) migrated to a position of around 52K. This is similar to the group consisting of mumps virus (MuV), Newcastle disease virus (NDV), human parainfluenza virus type 2 (PIV-2), human parainfluenza virus type 4 (PIV-4) and simian virus 5 (SV5), all with a P protein of approximately 45K to 55K. All these features indicate that LPMV is a member of the paramyxovirus genus.

The negative-stranded enveloped RNA viruses including the paramyxoviridae encode a membrane or matrix protein, which as well as being structural in the virion, has important functions during virus replication and maturation (Marx et al., 1974; Shimizu & Ishida, 1975; Yoshida et al., 1976; Büechi & Bäch, 1982; Peeples & Bratt, 1984).

The relationships between the members of the paramyxoviridae have been analysed by sequencing the viral genes and by serological means (Morrison, 1988; Ito et al., 1987). By comparing the amino acid sequence of the M proteins, three groups could be distinguished (Morrison, 1988; Limo & Yilma, 1990; Sheshberadaran & Lamb, 1990). The human parainfluenza virus type 3 (PIV-3), Sendai virus (SV), and the morbilliviruses are related by about 40%, with even higher homology within the morbilliviruses (Limo & Yilma, 1990). A second group consists of SV5, MuV and NDV, with NDV only distantly related to the others (Sheshberadaran & Lamb, 1990). Respiratory syncytial (RS) virus has a distinct M protein not related to the other groups. Thus the genus paramyxovirus could be divided into two groups on the basis of the sequence homology of the M protein: (i) SV and PIV-3, and (ii) MuV, SV5 and NDV. This is also in accordance with the studies of Ito et al. (1987), where the human paramyxoviruses are divided into two serologically distinct groups. Simian virus 41 was shown recently to be related serologically to the MuV, PIV-2, PIV-4 and SV5 group (Nishio et al., 1990).

The aim of this study was to produce evidence that LPMV is a paramyxovirus and to relate LPMV to the other members of the paramyxovirus genus. The complete nucleotide sequence of the M gene and the deduced amino acid sequence were determined and...
compared to other paramyxoviruses. A high level of amino acid identity to other paramyxovirus M proteins strongly confirms previous studies that LPMV is a member of the paramyxovirus genus.

Methods

Cells and viruses. Monolayers of porcine IBRS cells were infected with an isolate of LPMV (1984) that had previously been grown in cells for four passages. Virus was purified by repeated sucrose gradient centrifugation.

Generation and identification of cDNA clones. Viral RNA was extracted from virions digested by proteinase K (0.5 mg/ml) in 0.5% SDS and 10 mM-EDTA, with phenol-chloroform-isooamylalcohol. After precipitation the RNA was analysed by agarose gel electrophoresis.

Complementary DNA was synthesized using random hexanucleotides in a first-strand buffer containing 50 mM-Tris-HCl pH 8.3, 10 mM-MgCl2, 75 mM-KCl, 10 mM-DTT, 2 mM of each dNTP, 100 μg/ml bovine serum albumin (BSA) and 1000 units/ml of murine leukaemia virus reverse transcriptase. The second strand was synthesized in a second-strand buffer of 20 mM-Tris-HCl pH 7.5, 4 mM-MgCl2, 10 mM-(NH4)2SO4, 100 mM-KCl, 50 μg/ml BSA, 40 μM of each dNTP, 0.14 mM-NAD, 4.2 ng/ml DNA polymerase I and 8-8 units/ml RNase H. T4 DNA polymerase was then added to ensure that blunt-end double-stranded cDNA was obtained. The cDNA was ligated into pUC9 plasmids and transformed into E. coli DH5α.

Clones were screened by hybridization of radioactive labelled random-primed viral RNA.

DNA sequencing. To sequence the clones, nested unidirectional deletions were generated as described by Henikoff (1984). The alignment of sequences were done based on the algorithm of Needleman & Wunsch (1970). The deletions were selected, containing overlapping sequences, and DNA was sequenced according to standard procedures (Sanger et al., 1977) with M13 reverse and forward sequencing primers and T7 polymerase. To confirm the sequence, selected parts of the opposite strand were sequenced. Several LPMV-specific clones hybridizing to viral RNA were sequenced. In the mRNA sense L36 contained a long open reading frame (ORF) of 1185 nucleotides. Fifty-nine nucleotides upstream of the first ATG of this gene are underlined. The deduced amino acids are presented below the nucleotide sequence.

To confirm the sequence, selected parts of the opposite strand were sequenced. Several LPMV-specific clones hybridizing to viral RNA were sequenced. In the mRNA sense L36 contained a long open reading frame (ORF) of 1185 nucleotides. Fifty-nine nucleotides upstream of the first ATG of this gene are underlined. The deduced amino acids are presented below the nucleotide sequence.

Results

Nucleotide sequence of the M gene

Several LPMV-specific clones hybridizing to viral RNA and confirmed by Northern blot analysis (data not shown) were obtained. Two independently isolated overlapping clones designated L36 and L243 with inserts of approximately 2000 and 1000 nucleotides, respectively, were sequenced. In the mRNA sense L36 contained a long open reading frame (ORF) of 1185 nucleotides. Fifty-nine nucleotides upstream of the first ATG of this gene are underlined. The deduced amino acids are presented below the nucleotide sequence.

Fig. 1. Nucleotide sequence of the complementary DNA to viral RNA encoding the M protein of LPMV as determined from the clones L36 and L243. Poly(A) signals of the preceding gene (P gene) and of the M gene are underlined. The deduced amino acids are presented below the nucleotide sequence.
ORF a paramyxovirus putative poly(A) signal (AAAUUCUUUUUU, genomic sense) (Elango et al., 1988) was found. This probably corresponds to the poly(A) signal of the preceding P gene. The M protein-coding sequence from the first ATG to the termination codon contained 1107 nucleotides. Two potential poly(A) signals (AAAUCUUUUUUAAAG-CUAUAAUUUUU) were found 174 nucleotides downstream of the termination codon. Clone L243 confirmed the nucleotide sequence of the end of the P gene (data not shown), the poly(A) signal of the P gene and the first 650 nucleotides of the M protein gene. Starting from the first methionine in the ORF, a protein of 369 amino acids was deduced. The first methionine residue of the ORF is surrounded by signals corresponding to a eukaryotic ribosome-binding site (Kozak, 1986). The nucleotide sequence and deduced amino acid sequence are displayed in Fig. 1. The stop signals and ATGs of the sequence for the three reading frames are shown in Fig. 2. Only one of the reading frames is long enough to encode the M protein.

Properties of the M protein

The deduced protein consisted of 369 amino acids, with a calculated $M_r$ of 41,657. This is in good agreement with the $M_r$ of approximately 40,000 determined by SDS-PAGE (Sundqvist et al., 1990). The hydropathy profile of the M protein was determined and compared to those of MuV, SV5, NDV and canine distemper virus (CDV) (Fig. 3). The comparison revealed similar hydropathy
where eight, nine and seven of nine amino acids are identical in the corresponding sequences of MuV, SV5 and NDV respectively. The second domain is between residues 87 and 98 where seven (MuV), seven (SV5) and five (NDV) out of seven. Another typical feature of the M proteins of paramyxoviruses is the high number of basic residues, and the LPMV M protein has 14.6% basic residues as compared to +17.5, in the same range as calculated for NDV (+19.5), PIV-3 (+20.5) and MuV (+19.5).

The protein possesses an overall positive net charge of +17.5, in the same range as calculated for NDV (+19.5), PIV-3 (+20.5) and MuV (+19.5). Another typical feature of the M proteins of paramyxoviruses is the high number of basic residues, and the LPMV M protein has 14.6% basic residues as compared to 14.9% for the MuV M protein.

### Amino acid identity among paramyxovirus M proteins

By amino acid sequence alignment (Fig. 4) it was possible to recognize at least three highly conserved domains of the M protein. The first domain is between amino acid residues 57 and 65 of the LPMV M protein where eight, nine and seven of nine amino acids are identical in the corresponding sequences of MuV, SV5 and NDV respectively. The second domain is between residues 87 and 98 where seven (MuV), seven (SV5) and five (NDV) out of seven. Another typical feature of the M proteins of paramyxoviruses is the high number of basic residues, and the LPMV M protein has 14.6% basic residues as compared to 14.9% for the MuV M protein.

#### Discussion

The membrane or matrix proteins of paramyxoviruses have several important functions during infection. The M proteins are thought to mediate interactions between the glycoproteins in the plasma membrane, the ribonucleoprotein complex (RNP) and the host cell cytoskeleton during assembly and budding of the virus. In addition, the M protein was shown to be involved in viral transcription (Marx et al., 1974) and modification of the fusion protein function (Peeples & Bratt, 1984). The related vesicular stomatitis virus (VSV) M protein was shown recently to have at least two separated domains with different functions. One of these was responsible for the shut-off of cellular RNA synthesis and for the recognition of cellular factor(s) required for efficient RNA synthesis (Coulon et al., 1990). Lenard & Vanderoef.
(1990) demonstrated that the interaction of the VSV M protein with the virus bilayer lies within the first 19 amino acid residues of the N-terminal amphipathic segment. The M protein of paramyxoviruses may also have similar functional domains. The highly conserved regions of the M protein of LPMV, MuV, SV5 and NDV suggest that they have functional importance. Some residues which may be important in folding of the protein are also conserved (see Fig. 4).

In terms of M protein sequence homology, paramyxoviridae may be divided into three groups: (i) SV, PIV-3 and the morbiliviruses, (ii) MuV, SV5 and NDV and (iii) RS virus (Morrison, 1988; Sheshheradaran & Lamb, 1990; Limo & Yilma, 1990). Analysis of the nucleotide sequence and the deduced 369 amino acids of the M protein of LPMV and comparison of the amino acid sequence with that of other paramyxoviruses show a considerable degree of similarity to the paramyxoviruses MuV and SV5. Although we have not formally proven that the ORF and the predicted amino acids represent the M protein of LPMV, the resemblances to other M protein strongly suggest that this is the M protein gene. (Sheshheradaran & Lamb, 1990; Elliott et al., 1989; Limo & Yilma, 1990), and the hydrophatic profile of the protein strongly support that this is the M protein gene. Homology to the human MuV points to an evolutionary relationship between these viruses. The inter-relatedness between LPMV, MuV and SV5 of approximately 40% suggests a common ancestor of these viruses. In addition, sequence analysis of clones corresponding to other regions of the LPMV genome demonstrate a high relatedness towards the human PIV-2 and PIV-4 viruses (M. Berg et al., unpublished results).

LPMV causes an acute infection in piglets and the virus can be detected and recovered from various organs including the brain as early as the fourth and as late as the twentieth day post-infection, indicating a systemic infection (Stephano et al., 1988). The pronounced central nervous system (CNS) disturbances which follow LPMV infection are an unusual manifestation of paramyxovirus infection. Some other paramyxoviruses may affect the CNS as NDV does in chickens (Brandly, 1964), and CDV in dogs (Appel, 1987) and measles virus or MuV in rare cases after chronic infection of humans (Shaffer et al., 1942; Kilham, 1949). Defective M protein expression was initially suspected to be of selective advantage in persistent infection (Hall et al., 1979). Indeed, unusual hypermutation of the M gene may favour propagation of measles virus infections in the brain by conferring a selective advantage to the mutated genome (Cattaneo et al., 1988). However, defective M protein expression might not be the only viral determinant related to persistence. Other mechanisms may be likely in the case of LPMV. As a first step towards understanding the molecular basis of LPMV infection and pathogenesis, the viral genome was molecularly cloned. This first report presents the sequence of the M protein in comparison to other paramyxoviral M proteins.

Taken together the results of this study demonstrate that LPMV is most closely related to MuV and SV5. The high similarity of LPMV to these viruses indicates a close evolutionary relationship and provides evidence that LPMV should be classified as a genuine member of the genus paramyxovirus, in the same group as MuV, PIV-2, PIV-4 and SV5. In addition, clues to the functions of domains of the M protein can be found by comparing these related paramyxoviruses.

Thanks are due to Ann-Christin Bergvall for technical assistance. The work was supported by grants from the Swedish Agency for Research Cooperation (SAREC) and the Swedish Council for Forestry and Agricultural Research.

References


Paramyxoviruses M protein gene sequence 1049


(Received 15 November 1990; Accepted 17 January 1991)